Review

CellPress

Special Issue: Manifesting Synthetic Biology

Synthetic protein switches: design principles and applications

Viktor Stein and Kirill Alexandrov

The University of Queensland, Institute for Molecular Biosciences, Brisbane, QLD 4072, Australia

Protein switches are ubiquitous in biological signal transduction systems, enabling cells to sense and respond to a variety of molecular queues in a rapid, specific, and integrated fashion. Analogously, tailor-engineered protein switches with custom input and output functions have become invaluable research tools for reporting on distinct physiological states and actuating molecular functions in real time and *in situ*. Here, we analyze recent progress in constructing protein-based switches while assessing their potential in the assembly of defined signaling motifs. We anticipate such systems will ultimately pave the way towards a new generation of molecular diagnostics and facilitate the construction of artificial signaling systems that operate in parallel to the signaling machinery of a host cell for applications in synthetic biology.

Synthetic biology of signal transduction

The rational construction of artificial signaling systems is a key goal of synthetic biology. This encompasses all levels of complexity, ranging from proteins to pathways, networks, and, ultimately, organisms, and has application for molecular diagnostics, cell-based biosensors, therapeutics, and industrial biotechnology [1,2]. In addition, a capacity to engineer biological signaling systems with predictable behavior provides ultimate proof to scientific models describing biological processes [1].

Constructing artificial signaling systems has been realized predominantly with synthetic gene circuits, in which rational engineering strategies are supported by the modular organization and function of transcription factors and their DNA response elements [3,4]. Similarly, aptamers and ribozymes have been recombined to create functional nucleic acids that can sense and amplify distinct molecular cues [5] or exert post-transcriptional control on gene expression [6]. However, the limited chemical diversity of nucleic acids compared with amino acids ultimately limits their functionality. Furthermore, transcription-based signaling circuits are inherently slow, with typical response times on the scale of hours [7]. By contrast, protein-based signaling circuits operate orders of magnitude faster and feature diverse enzymatic outputs [7]. However, engineering such systems has proven challenging, especially at the molecular level, where signaling is controlled by distinct protein switches. These can either be based on allosterically regulated proteins that couple input to output solely through conformational changes or be composed of modular receptors, transducers, and actuators that process molecular cues in a concerted fashion through the induced colocalization of distinct signaling components. The individual designs can range from highly integrated based on structurally intertwined receptors and actuators, where conformational changes are transmitted through networks of residues that are adjacent in tertiary, but not necessarily in primary structure (Figure 1A), to modular, where conformational changes are limited to the linker regions that separate functionally and structurally distinct receptor and actuator domains (Figure 1B), to highly modular, where signaling cues are transmitted through the induced colocalization of molecularly distinct signaling components (Figure 1C).

Applications for protein switches are numerous. In diagnostics, protein switches can detect analytes as components of inexpensive homogeneous assays that do not require specialized equipment or time-consuming incubation and washing steps characteristic of immunoassays [8]. Protein switches have also proven invaluable in the quantitative imaging of molecular processes in cells [9]. In addition, protein switches can control the activity of key signaling proteins by non-invasive means, such as with light or biochemically inert ligands, which act orders of magnitude faster than inducible gene expression-based systems.

Here, we review recent progress in the construction of protein-based switches for monitoring and actuating molecular and cellular functions while identifying aspects critical for their successful design. Overall, this should greatly facilitate the challenging task of constructing protein-based switches, which has so far proven intractable to computational design methods [10,11] (Box 1).

Allosteric fluorescent protein switches

Genetically encoded Ca^{2+} sensors constitute the first generation of protein switches that exploited fluorescent proteins (FPs, see Glossary) for generating a measureable read-out [12,13]. Allosteric binding receptors for Ca^{2+} were generated by fusing calmodulin (CaM) to a CaM-binding peptide (CaM-BP) derived from the myosin light chain kinase. As CaM binds Ca^{2+}, CaM-BP associates with CaM, causing the receptor to transition from an extended

Corresponding authors: Stein, V. (v.stein@imb.uq.edu.au); Alexandrov, K. (k.alexandrov@uq.edu.au).

Keywords: synthetic biology; protein switches; protein engineering; molecular diagnostics; signaling; biosensors.

0167-7799/

 http://dx.doi.org/10.1016/j.tibtech.2014.11.010

Glossary

Actuator: protein modules that undergo changes in easily detectable biophysical properties, including binding, catalytic activity, or other type of signaling chemistry, such as fluorescence or bioluminescence.
Affinity clamps: a class of artificially engineered allosteric binding receptors where a primary protein-binding domain is fused to a secondary enhancer domain that specifically recognizes the ligand-bound conformation of the primary binding domain resulting in the formation of a sandwich complex.
Allosteric binding receptor: refers to protein-based receptors that undergo conformational changes upon binding of their target ligand.
Alternative frame folding (AFF): refers to a generally applicable procedure for creating allosteric binding receptors and actuators based on competing, partially folded protein fragments that can be selectively stabilized in response to specific molecular queues.
Autoinhibitory domains (AI domains): directly bind the active or allosteric site of an enzyme to inhibit catalytic or binding functions.
***Avena sativa* phototropin 1 Lov2 domain (AsLov2):** refers to a flavin mononucleotide-containing protein domain that undergoes a reversible conformational change upon illumination, which is exploited by several light-responsive synthetic protein switches.
β-Lactamase (BLA): a model reporter enzyme that confers resistance to β-lactam antibiotics, the activity of which can be readily detected using both colorimetric and fluorescent read-outs.
β-Lactamase inhibitory protein (BLIP): a competitive, active site-directed inhibitor of BLA that is used as an AI domain to create allosterically regulated BLA-based protein switches.
Bioluminescent resonance energy transfer (BRET): refers to the distance-dependent energy transfer between a bioluminescent donor and a fluorescent acceptor. BRET-dependent read-outs are frequently used to resolve bimolecular interaction events or conformational changes in allosterically regulated protein switches with luciferase and FP-tagged binding receptors.
Bottom-up design: the assembly of protein switches, signaling motifs, pathways, and networks from well-characterized component parts using rational engineering principles.
Calmodulin (CaMs): a family of calcium-dependent allosteric binding receptors that transition from an extended to a compact conformation upon binding CaM-BPs.
CaM-binding peptides (CaM-BP): a family of peptides that specifically bind to CaM.
Circular protein permutation: a protein engineering procedure where the native N and C termini are fused and new N and C termini are created at a different site, leaving the overall structure intact.
Dihydrofolate reductase (DHFR): a model reporter enzyme that catalyzes the conversion of dihydrofolic acid to tetrahydrofolic acid using NADPH as an electron donor.
Fibronectin type III (FN3) domain: a class of immunoglobulin-like protein domains that, similar to antibody fragments, can be tailor-engineered to recognize a variety of protein-based targets.
Firefly luciferase (FFL): a model reporter enzyme that catalyzes the oxidation of D-luciferin (LH2) to oxyluciferin with the concomitant emission of light.
FK506-binding protein (FKBP12): forms part of a model protein–protein interaction module that binds FRB in a rapamycin-dependent manner.
FKBP rapamycin-binding protein (FRB): forms part of a model protein–protein interaction module that binds FKBP12 in a rapamycin-dependent manner.
Fluorescent proteins (FP): proteins that fluoresce upon illumination with light and are frequently used to visualize bimolecular interaction events or conformational changes in allosteric protein receptors by means of distance-dependent FRET between two FPs.
Fluorescence resonance energy transfer (FRET): refers to the distance-dependent energy transfer between a donor and acceptor fluorophore. FRET-dependent read-outs are frequently used to resolve bimolecular interaction events or conformational changes in allosterically regulated protein switches with FP-tagged binding receptors.
Guanine nucleotide exchange factors (GEFs): key regulatory proteins that mediate the activation of small GTPases as they catalyze the exchange of GDP for GTP.
Luciferase-based indicators of drugs (LUCIDs): a class of semisynthetic small molecule sensors where binding of a target analyte triggers conformational changes in the sensor that is subsequently detected by BRET.
Maltose-binding protein (MBP): an allosteric protein binder belonging to the periplasmic binding protein family that undergoes a rigid domain movement upon binding maltose. In association with BLA, MBP has been used to construct various allosteric protein switches by means of domain insertion.
Mitogen-activated protein kinases (MAPK): a highly conserved, eukaryotic signaling system that transmits signals from receptors at the plasma membrane through a phosphorylation cascade to the nucleus.
Protein fragment complementation assays (PCAs): refer to bimolecular protein–protein interaction assays that transduce a binding event through the assembly of a fully functional reporter enzyme.
PSD95-Dlg1-Zo1 (PDZ): a family of protein-binding domains that specifically bind distinct C-terminal peptide motifs.
Rapamycin: refers to a small molecule drug that mediates association between FKBP12 and FRB.
SNAP-Tag: a site-specific protein conjugation system based on engineered mutants of human alkyl guanine transferase (hAGT) that mediate the irreversible conjugation of SNAP-tag fusion proteins with benzyl-guanine-labeled small molecule tags that include fluorophores, affinity handles, and small molecule drugs.
SNAP-tag based indicator with a fluorescent intramolecular tether (SNIFIT): a class of semisynthetic small molecule sensors where binding of a target analyte triggers conformational changes in the sensor that is subsequently detected by FRET.
Src homology 2 (SH2) domains: a family of protein domains that specifically bind phosphopeptide motifs.
Src homology 3 (SH3) domains: a family of protein domains that specifically bind proline-rich peptide motifs.
Top-down design: the construction of protein-based switches, signaling motifs, pathways, and networks through the step-wise modification of existing systems.
Transducer: generically refers to protein-based binders, enzymes, modular linker elements, and networks of interacting residues that translate a conformational change in the binding of a receptor to an actuator.

to a compact state. In the original design, this conformational change was detected through distance-dependent changes in the efficiency of fluorescence resonance energy transfer (FRET) between a yellow- (YFP) and cyan-fluorescent protein (CFP) [12]. In a more integrated design, the Ca^{2+} receptor was inserted internally into GFP, resulting in Ca^{2+}-triggered modulation of GFP fluorescence intensity [13,14]. Iterative improvements over the course of 15 years have yielded highly sensitive calcium sensors that can detect the comparatively small calcium bursts that occur during axon potential firing in neurons [15,16].

Based on these initial blueprints, more than 100 intramolecular FRET sensors have been developed to monitor a variety of molecular queues, ranging from protein–peptide, antibody–epitope, and protein–small molecule interactions to the detection of proteolysis, post-translational modifications, and mechanical properties, such as tension at cell–cell junctions [17–19]. Biophysical studies demonstrated that intramolecular FRET sensors with the largest signal change exploit the natural propensity of FPs to dimerize, which enables them to toggle between two defined molecular states in the presence and absence of a molecular queue. This sets them apart from FPs that do not form such interactions and exist in poorly defined conformational and spectral ensembles [17]. In practice, this requires the construction of linkers that are sufficiently rigid to counteract the dimerization of FPs in the absence of a molecular queue, but sufficiently loose to ensure efficient FRET and folding of the fluorophores.

Allosteric fluorescent protein switches based on mutually exclusive binding interactions

Given that only a few naturally occurring protein families exhibit conformational changes sufficiently large for sensor construction, allosteric binding receptors frequently have to be engineered. Here, modular design strategies that utilize generic receptor modules and minimize the need to engineer protein-based binders *de novo* are preferred.

Affinity clamps

In 'affinity clamping', two binders are connected through a flexible linker, which enables them to form a sandwich complex around a target ligand. Proof-of-concept was